

IN THE CLAIMS:

The following listing of claims replaces all prior claim versions and listings:

Claims 1-60 (cancelled)

~~Claim 61.~~ (New) A method for retrieving answers to questions from an information retrieval system comprising:

✓ classifying questions into a plurality of question types;
✓ generating candidate query transformations for each question type from a collection of question-answer pairs, where the candidate query transformation for a question type is automatically derived from the answers associated with the questions in the question-answer pairs which match the question type;

answer
is
derived
from
a
document

✓ evaluating performance of the generated candidate query transformations by executing the information retrieval system using queries formulated by applying the candidate query transformations to the questions in the question-answer pairs and comparing results from the information retrieval system to the answers in the question-answer pairs; and

running by
computer
results (answers)
on a
document

✓ selecting for the information retrieval system one or more candidate query transformations for each question type based on the evaluation of the performance of the candidate query transformations.

~~Claim~~ 62. (New) The method of claim 61, wherein different candidate query transformations can be generated if a different collection of question-answer pairs is used and wherein different candidate query transformations can be selected if a different information retrieval system is executed.

~~Claim~~ 63. (New) The method of claim 61, wherein a candidate query transformation for a question type replaces one or more question words in a query for the information retrieval system with one or more answer words, the question words being defined by the question type and the answer words being automatically derived from the answers associated with the questions in the question-answer pairs which match the question type.

~~Claim~~ 64. (New) The method of claim 61, wherein computational requirements for evaluating the performance of the candidate query transformations are reduced by filtering the candidate query transformations before executing the information retrieval system.

~~Claim~~ 65. (New) The method of claim 64, wherein the candidate query transformations are filtered by using term weights estimating selectivity of the candidate query transformation with respect to a question type.

~~Claim~~ 66. (New) The method of claim 65, wherein the term weights are applied to phrases as well as single word terms used as candidate query transformations.

~~Claim 67.~~ (New) The method of claim 65 wherein the term weights are computed by

$$w_i^{(1)} = \frac{(r + 0.5)/(R - r + 0.5)}{(n - r + 0.5)/(N - n - R + r + 0.5)}$$

where r is a number of relevant answers containing the candidate query transformation in the collection of question-answer pairs, N is a number of question-answer pairs in the collection, R is a number of relevant answers in the collection, and n is a number of answers in the collection containing the candidate query transformation.

~~Claim 68.~~ (New) The method of claim 65 wherein the filtering of the candidate query transformations chooses candidate query transformations with a highest selection weight, the selection weights computed for each candidate query transformation based on co-occurrence statistics and on the term weights.

(Col. 31, 1-10)
Col. 34
(1-9)

~~Claim 69.~~ (New) The method of claim 68 wherein the selection weights and the term weights are applied to phrases as well as single word terms used as candidate query transformations.

~~Claim 70.~~ (New) The method of claim 69 wherein the filtering of the candidate query transformations includes sorting the candidate query transformations into buckets according to a number of words in a phrase in a candidate query transformation.

~~Claim~~ 71. (New) The method of claim 64 wherein the candidate query transformations are filtered by discarding transformations which include content words.

~~Claim~~ 72. (New) The method of claim 61 wherein the results from the information retrieval system are compared to the answers in the question-answer pairs by generating sub-documents from documents retrieved by the information retrieval system using the queries formulated by applying the candidate query transformations and by comparing the sub-documents to the answers in the question-answer pairs.

~~Claim~~ 73. (New) The method of claim 72 wherein the sub-documents are generated by extracting overlapping portions of the documents starting at successive positions in the documents.

~~Claim~~ 74. (New) The method of claim 72 wherein a similarity score assigned to a document with respect to an answer is computed by finding a sub-document in the document with a highest similarity score to the answer and assigning the highest similarity score to the document as a whole.

~~Claim~~ 75. (New) The method of claim 61 wherein a plurality of candidate query transformations are selected for each question type for the information retrieval system, the results using the plurality of candidate query transformations weighted using weights derived from the evaluation of the performance of the candidate query transformations.

~~Claim 76.~~ (New) The method of claim 75 wherein the weights for the plurality of candidate query transformations for a question type are based on an average similarity score for documents retrieved using the candidate query transformations with the information retrieval system.

Pedersen et al
5,442,778

~~Claim 77.~~ (New) The method of claim 61, wherein the question types used to classify the questions are automatically generated from the questions in the collection of question-answer pairs.

~~Claim 78.~~ (New) The method of claim 77 wherein the question types are generated by computing a frequency of n -grams anchored at a beginning of the questions in the collection of question-answer pairs.

~~Claim 79.~~ (New) A method for retrieving documents from an information retrieval system, the method comprising:

receiving a query;

classifying the query in accordance with one of a plurality of question types;

applying a query transformation to the query based on the question type, resulting in a transformed query which is processed by the information retrieval system, where the query transformation has been selected for the question type for the information retrieval system based on

generating candidate query transformations for each question type from a collection of question-answer pairs, where the candidate query transformation for a question type

were automatically derived from the answers associated with the questions in the question-answer pairs which match the question type; and on

evaluating performance of the generated candidate query transformations by executing the information retrieval system using queries formulated by applying the candidate query transformations to the questions in the question-answer pairs and comparing results from the information retrieval system to the answers in the question-answer pairs.

~~Claim 80.~~ (New) The method of claim 79, wherein the query transformation for the question type creates the transformed query by replacing one or more question words in the query for the information retrieval system with one or more answer words, the question words being defined by the question type and the answer words being automatically derived from the answers associated with the questions in the question-answer pairs which match the question type.

~~Claim 81.~~ (New) The method of claim 80 wherein the answer words used in the query transformation are treated by the information retrieval system as a single phrase.

~~Claim 82.~~ (New) The method of claim 79, wherein more than one query transformation is applied to the query, resulting in a plurality of transformed queries which are each processed by the information retrieval system and wherein results from the information retrieval system using each transformed query are combined to form combined results for the query.

~~Claim 83.~~ (New) The method of claim 82, wherein each query transformation has a weight associated with it and wherein the results from the information retrieval system using each transformed query are weighted using the associated weight when combined to form the combined results for the query.

~~Claim 84.~~ (New) The method of claim 83 wherein the combined results from the information retrieval system are ranked by generating sub-documents from documents retrieved by the information retrieval system using the transformed queries and by comparing the sub-documents to the transformed queries.

~~Claim 85.~~ (New) The method of claim 84 wherein the sub-documents are generated by extracting overlapping portions of the documents starting at successive positions in the documents.

~~Claim 86.~~ (New) The method of claim 84 wherein a similarity score assigned to a document with respect to the query is computed by finding a sub-document in the document with a highest similarity score to the transformed query and assigning the highest similarity score to the document as a whole.

~~Claim 87.~~ (New) The method of claim 86 wherein the weights for the plurality of query transformations for a question type are applied to an average similarity score for documents retrieved using the transformed query with the information retrieval system.